



Supplement of

Constraints on NO_x emission in Thailand using GEMS satellite data

Worapop Thongsame et al.

Correspondence to: Worapop Thongsame (t.worapop@gmail.com)

The copyright of individual parts of the supplement might differ from the article licence.

S1 The impact of the hyperparameter on IFDMB

In this section, we evaluate the impact of C_1 and C_2 parameters on the IFDMB performance. IFDMB with fixed diurnal update is applied to four different sets of $[C_1, C_2]$ parameters, and we evaluate IFDMB solution as a function of iteration for each. The four $[C_1, C_2]$ parameters are $[2 \times 10^{15}, 1]$, $[2 \times 10^{15}, 10]$, $[1 \times 10^{14}, 1]$, and $[1 \times 10^{14}, 10]$. We find that iteration five is the optimal iteration for $[2 \times 10^{15}, 1]$ and $[2 \times 10^{15}, 10]$, and iteration two is optimal for $[1 \times 10^{14}, 1]$ and $[1 \times 10^{14}, 10]$. All four parameters produce similar spatial patterns in model performance (Table S1). The IFDMB with fixed diurnal update significantly reduces both NMB and NRMSE relative to GEMS observations across Thailand, particularly in the Bangkok Metropolitan Region.

Increasing C_1 from 1×10^{14} to 2×10^{15} molecules cm^{-2} suppresses the influence of satellite constraints, resulting in higher NMB and NRMSE values for the $[2 \times 10^{15}, 1]$ and $[2 \times 10^{15}, 10]$ cases, especially in Central and East Thailand (Figs. S1 and S2). This is attributed to smaller emission updates under high C_1 . Even though the σ_o in these regions is low (on the order of $0.5 - 0.2 \times 10^{13}$ molecules cm^{-2}), the corresponding NO_2 VCDs are also relatively low, which leads to a small $J_{\text{prediction}}$ compared to the $J_{\text{parameter}}$. This causes the solution to remain close to the prior emissions. However, a larger C_1 value helps prevent overfitting in rural North Thailand, where σ_o is much higher.

Increasing C_2 from 1 to 10 $\text{mol km}^{-2} \text{hr}^{-1}$ has minimal impact in the $[1 \times 10^{14}, 1]$ and $[1 \times 10^{14}, 10]$ cases. While the percentage change in emissions is relatively larger in $[1 \times 10^{14}, 10]$, the improvements in NMB and NRMSE are comparable: $[1 \times 10^{14}, 1]$ reduces NMB by 66% and NRMSE by 62%, whereas $[1 \times 10^{14}, 10]$ achieves 47% and 48% reductions, respectively (Table S1). This is because the larger emission adjustments occur in low-emission areas, thus contributing little to the overall NO_2 VCD improvements. In contrast, the effect of increasing C_2 is more pronounced in the $[2 \times 10^{15}, 1]$ and $[2 \times 10^{15}, 10]$ cases. A higher C_2 value (10 $\text{mol km}^{-2} \text{hr}^{-1}$) allows emissions in Central and East Thailand to be more effectively constrained by GEMS data compared to $C_2 = 1$ (Figs. S1 and S2). Consequently, NMB and NRMSE improve significantly in $[2 \times 10^{15}, 1]$ with $\sim 30\%$ reduction. The $[2 \times 10^{15}, 10]$ yields reductions of approximately 47% in both metrics.

The parameter C_1 notably influences the relative errors between the Bangkok Metropolitan Region and North Thailand. A low C_1 value (1×10^{14} molecules cm^{-2}) permits large adjustments even in high error areas (Fig. S3) in GEMS like North Thailand. This greatly decreases NMB and RMSE in the Bangkok Metropolitan Region, but also excessively increases bias and error in North Thailand (Figs. S1 and S2). In contrast, a high C_1 (2×10^{15} molecules cm^{-2}) constrains adjustments in the North, preventing over-correction and lowering bias and error in that region. However, it also limits the improvement in the Bangkok Metropolitan Region.

The second parameter, C_2 , primarily affects adjustments for smaller emission sources (such as rural areas). Using a smaller C_2 limits the changes in rural emissions (Fig. S3), keeping those values closer to the prior and yielding lower bias and error against TROPOMI. In contrast, a larger C_2 allows greater adjustments for low-emission areas, which led to higher bias and error

in rural North due to overfitting to the GEMS data. Overall, the adjustable parameters $[C_1, C_2]$ in IFDMB offer flexibility to manage the mismatch between GEMS and TROPOMI in the North by moderating the influence of emission updates in high-error regions and balancing between prior and posterior estimates.

Table S1: Monthly statistical evaluation between WRF-Chem and GEMS NO₂ VCD of IFDMB for each $[C_1, C_2]$ parameters: $[1 \times 10^{14}, 1]$, $[2 \times 10^{15}, 1]$, $[1 \times 10^{14}, 10]$, and $[2 \times 10^{15}, 10]$.

	$[1 \times 10^{14}, 1]$			$[2 \times 10^{15}, 1]$			$[1 \times 10^{14}, 10]$			$[2 \times 10^{15}, 10]$		
	NMB	NRMSE	<i>R</i>	NMB	NRMSE	<i>R</i>	NMB	NRMSE	<i>R</i>	NMB	NRMSE	<i>R</i>
All hour	28.78	58.01	0.83	58.95	106.51	0.67	31.90	56.25	0.84	44.12	78.38	0.80
8 AM	182.35	238.83	0.74	239.49	320.79	0.56	184.21	235.57	0.75	207.88	266.08	0.72
9 AM	137.07	186.46	0.81	191.50	268.58	0.69	142.48	187.51	0.80	163.54	215.74	0.81
10 AM	68.16	116.39	0.80	109.21	179.89	0.69	74.24	117.91	0.78	89.83	142.27	0.79
11 AM	22.23	67.94	0.77	51.59	112.94	0.63	26.71	67.61	0.77	38.56	87.41	0.75
12 PM	-3.62	50.73	0.72	19.85	84.19	0.55	-0.24	48.86	0.73	9.10	66.10	0.69
2 PM	-6.81	56.53	0.65	16.52	90.37	0.48	-4.75	53.50	0.68	5.78	73.10	0.61
3 PM	2.28	59.95	0.66	27.87	98.65	0.51	5.07	58.27	0.69	15.69	78.92	0.63

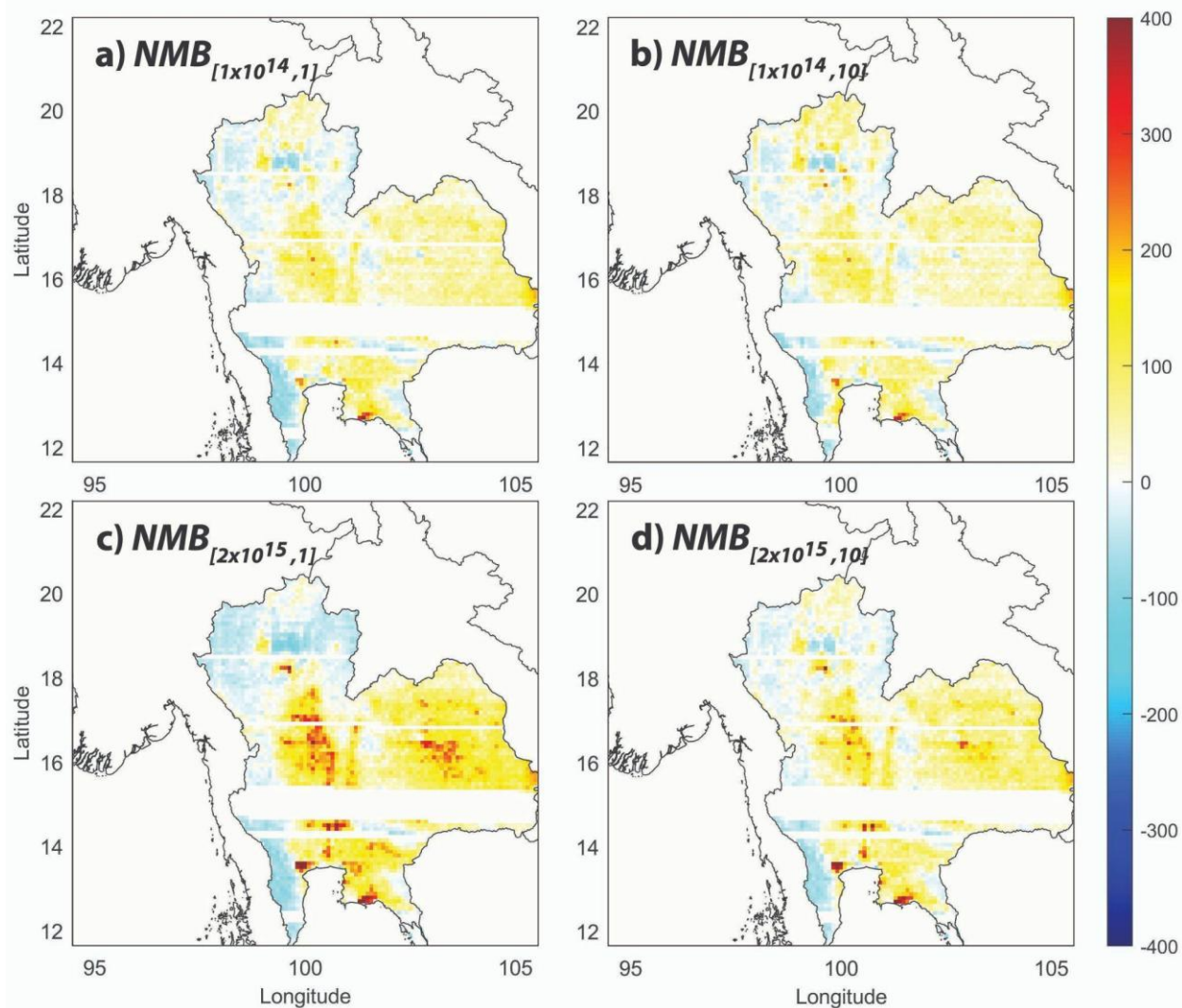


Figure S1: Normalized Mean Bias (NMB) of hourly NO_2 VCD between WRF-Chem and GEMS of IFDMB for September 2023 using $[C_1, C_2]$: (a) $[1 \times 10^{14}, 1]$, (b) $[1 \times 10^{14}, 10]$, (c) $[2 \times 10^{15}, 1]$, and (d) $[2 \times 10^{15}, 10]$.

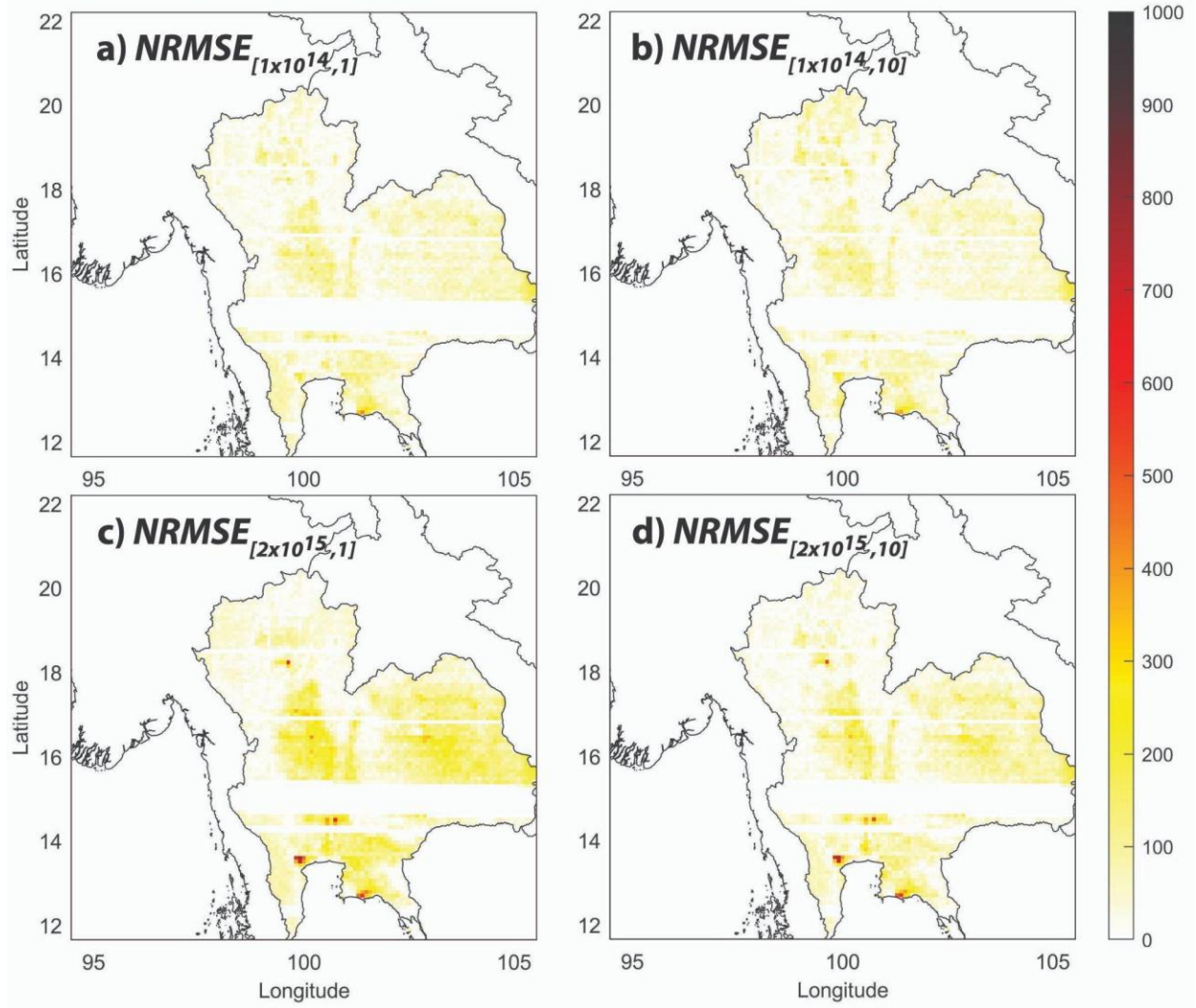


Figure S2: Normalized Root Mean Square (NRMSE) of hourly NO_2 VCD between WRF-Chem and GEMS of IFDMB for September 2023 using $[C_1, C_2]$: (a) $[1 \times 10^{14}, 1]$, (b) $[1 \times 10^{14}, 10]$, (c) $[2 \times 10^{15}, 1]$, and (d) $[2 \times 10^{15}, 10]$.

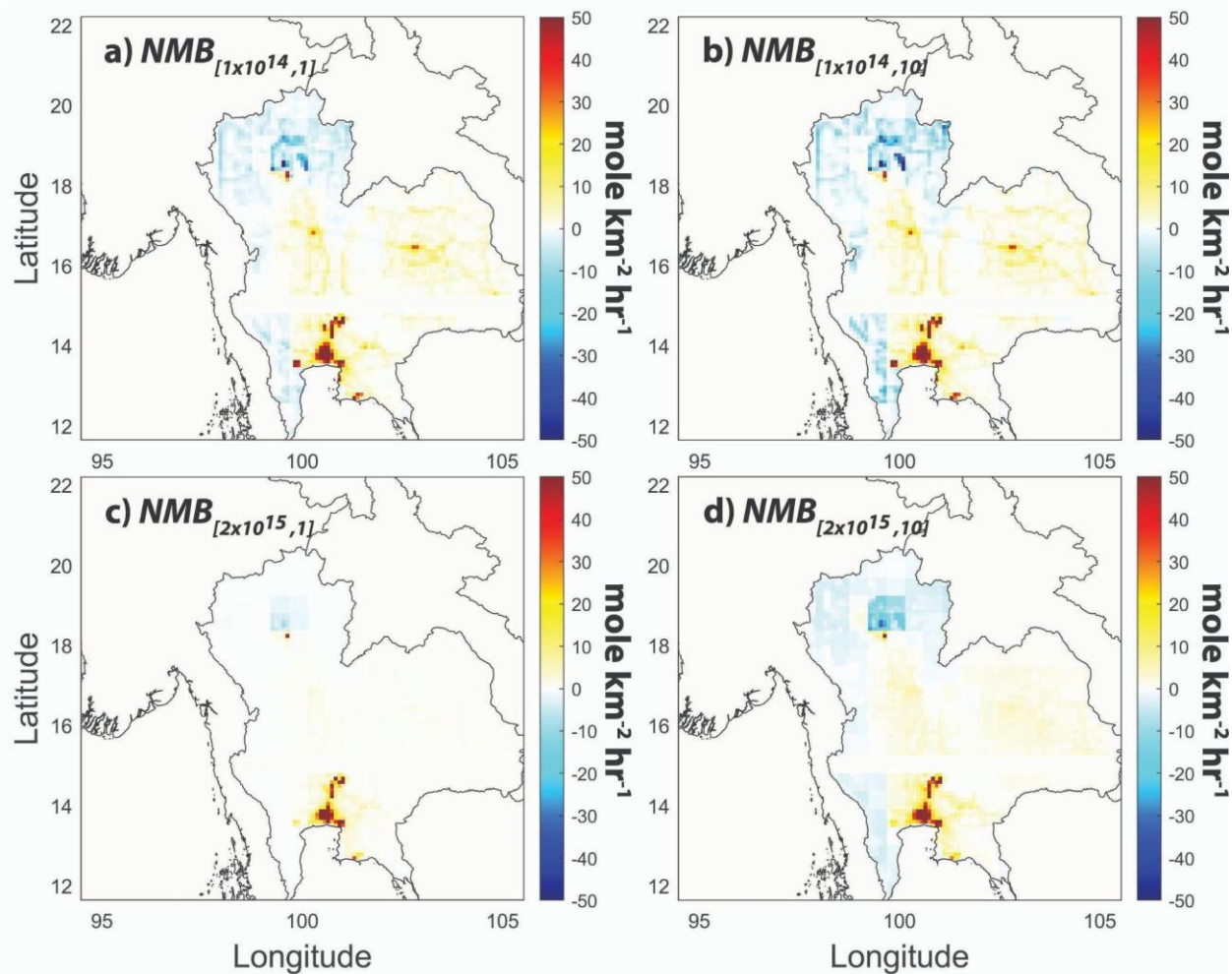


Figure S3: The emission changes from prior for IFDMB with different $[C_1, C_2]$ values: (a) changes from $[1 \times 10^{14}, 1]$, (b) changes from $[1 \times 10^{14}, 10]$, (c) changes from $[2 \times 10^{15}, 1]$, and (d) changes from $[2 \times 10^{15}, 10]$.

S2 The comparison between GEMS DVCF and GEMS DOAS

The retrieval algorithm impacts the NO₂ VCDs from the GEMS satellite. Yang et al. (2023) use the Direct Vertical Column Fitting (DVCF) algorithm, whereas GEMS version 3 relies on the DOAS algorithm. The two products also differ in their prior models. The DVCF algorithm (Huang et al., 2022; Yang et al., 2014) finds the modeled radiance spectra (I_{TOA}) that match the satellite measured spectra (I_m) by minimizing the cost function as shown in Eq (S1).

$$\ln I_m(\lambda) - \ln I_{TOA}(\lambda) = V \int_0^\infty \frac{\partial \ln I_{TOA}(\lambda)}{\partial \tau_z} S_z \sigma(\lambda, T_z) dz - \sum_{i=1}^m \xi_i \sigma_i(\lambda, T_i) + \sum_{k=0}^{n-1} \frac{\partial \ln I_{TOA}(\lambda)}{\partial R} \Delta R_k (\lambda - \lambda_0)^k + \varepsilon \quad (S1)$$

where λ is the wavelength and V is the NO₂ vertical column. S_z is the shape factor derived from the normalized vertical profile at the altitude (z). A more detailed explanation of the algorithm is in Yang et al. (2014).

The prior model used in GEMS DVCF is GEOS-CF using the HTAP inventory, while that in GEMS DOAS is GEOS-Chem using the ASIA-AQv3 inventory. These differences contribute to variations in the slant columns, vertical columns, and averaging kernels between GEMS DVCF and GEMS DOAS.

Because the averaging kernels differ between GEMS DVCF and GEMS DOAS, direct comparison is not feasible; therefore, we assess the relative NO₂ distributions instead. GEMS DVCF has relatively higher NO₂ over the Bangkok Metropolitan Region compared to North Thailand, while GEMS DOAS has higher NO₂ levels in the North (Fig. S4). Notably, GEMS DOAS has a large NO₂ signal over Lampang due to the high NO_x emissions in ASIA-AQv3. This high NO₂ is absent in both GEMS DVCF and TROPOMI. The uncertainties in ASIA-AQv3 lead to high retrieval error in GEMS DOAS in Lampang. Furthermore, both GEMS ACX and GEMS v3 exhibit artificial striping, implying that the stripes originate from the instrument itself rather than the retrieval algorithm.

We also compare the magnitude of NO₂ columns between the two retrievals using the baseline WRF-Chem simulation as a reference. When compared with GEMS DOAS, the baseline simulation generally overestimates NO₂ VCDs across Thailand, with an NMB of 84 and an NRMSE of 152 (Fig. S5). The model notably overestimates NO₂ in the Bangkok Metropolitan Region but underestimates it in North and West Thailand. For GEMS DVCF, we normalize the bias and error using the retrieval error from GEMS DOAS, as there is no error provided in GEMS DVCF. The baseline simulation overestimates NO₂ VCDs across most of the domain, with an NMB of 198 and NRMSE of 216, except in the western Bangkok Metropolitan Region (Fig. S5). This suggests that the CAMS-GLOB-ANT inventory overestimates NO_x emissions in Thailand when compared with GEMS DOAS. In contrast, CAMS-GLOB-ANT tends to underestimate NO_x emissions in North Thailand relative to GEMS DOAS. Although GEMS DVCF shows good agreement with CAMS-GLOB-ANT over major NO_x sources such as Bangkok, Rayong, and the Mae Moh power plant, it yields much lower NO₂ VCDs in small cities

and background regions compared to these urban and industrial areas. Additionally, GEMS DVCF shows a broader spatial spread of NO₂ over the Bangkok Metropolitan Region than TROPOMI. Considering these characteristics, we adopt the GEMS DOAS product for this study.

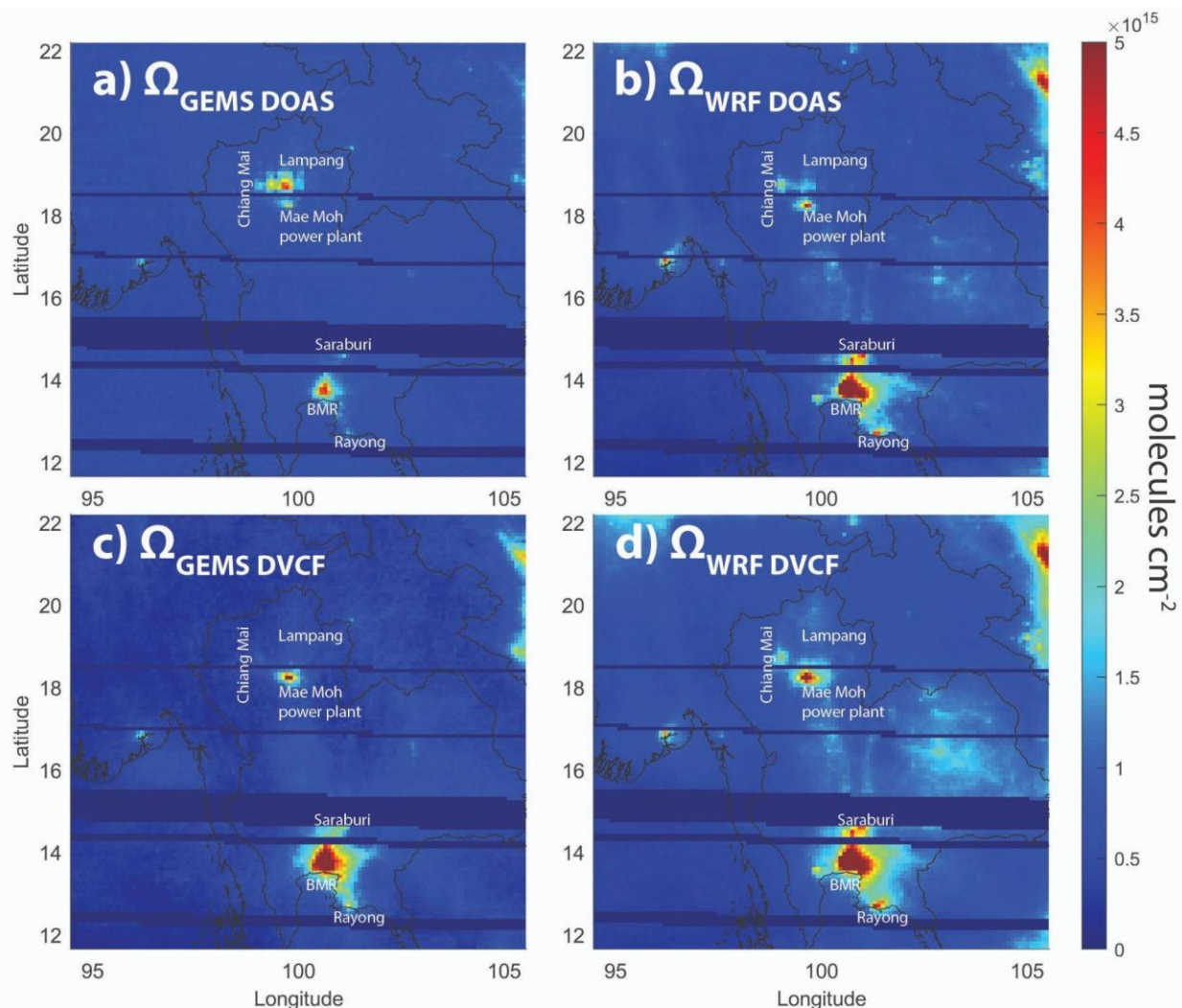


Figure S4: Monthly average tropospheric NO₂ VCD in September 2023 from GEMS DOAS (a) and GEMS DVCF (c), the baseline simulation applying the averaging kernel from GEMS DOAS (b) and GEMS DVCF (d). High NO₂ VCDs are observed over urban areas such as Chiang Mai, Lampang, and the Bangkok Metropolitan Region (BMR), as well as over the Mae Moh power plant and industrial zones in Saraburi and Rayong.

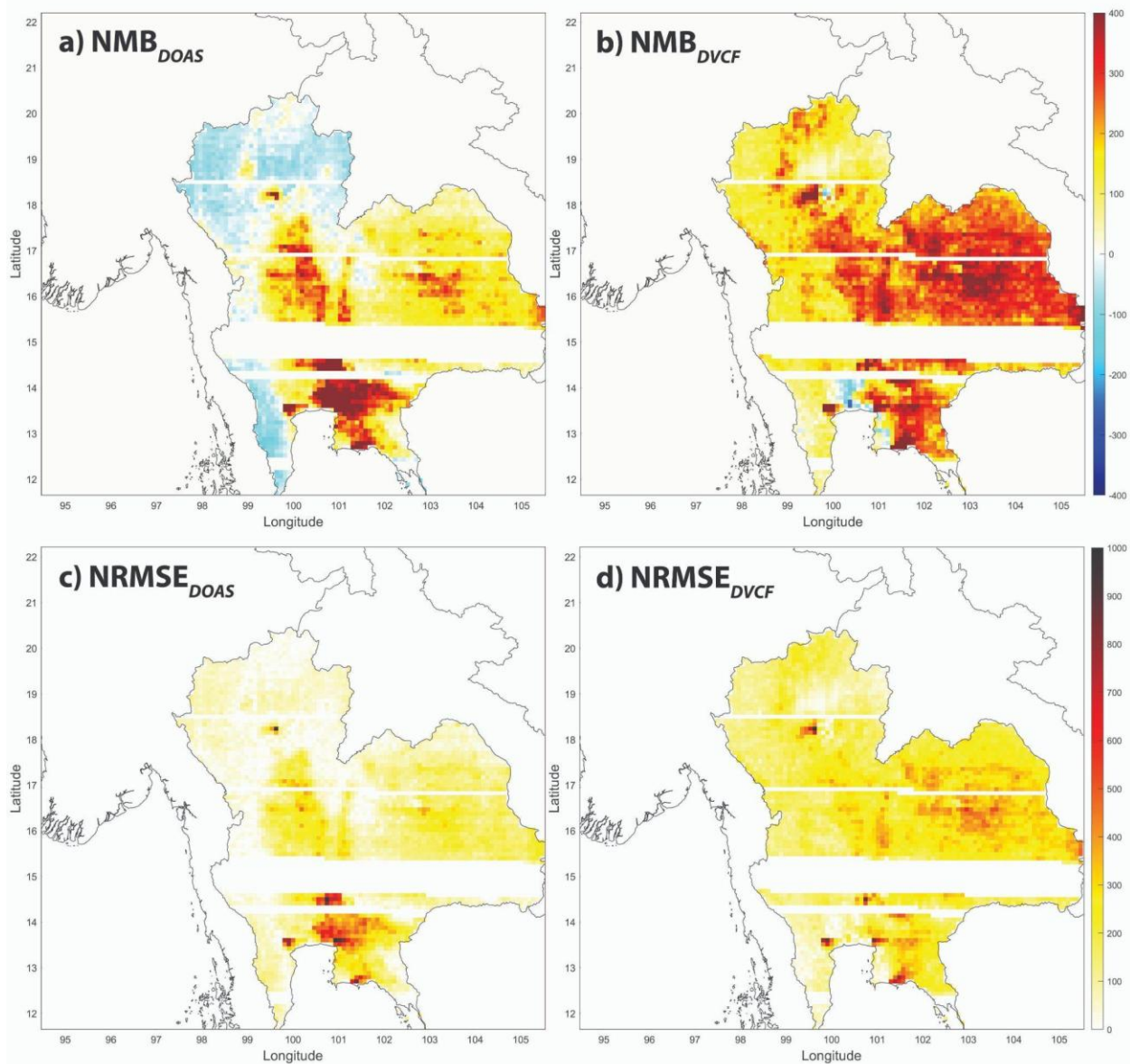


Figure S5: Normalized Mean Bias (NMB) of NO_2 VCD between the baseline simulation and GEMS DOAS (a) and GEMS DVCF (b). Normalized Root Mean Square Error (NRMSE) of NO_2 VCD between the baseline simulation and GEMS DOAS (c) and GEMS DVCF (d).

S3 Uncertainty in GEMS DOAS due to ASIA-AQv3 Emission Inventory

According to the GEMS DOAS team (private communication), the operational GEMS version 3 uses the ASIA-AQv3 anthropogenic emissions inventory as input for its prior GEOS-Chem model. This prior is used to derive key retrieval parameters, including air mass factors, shape factors, and averaging kernels. In our study period (September 2023), biomass burning in Thailand is minimal (late wet season), so anthropogenic emissions are the dominant source of NO_2 . Consequently, the ASIA-AQv3 inventory can exert a significant influence on the retrieved GEMS NO_2 VCDs over Thailand during this time.

We discovered that the ASIA-AQv3 inventory contains an anomalously high NO_x point source in Lampang. The magnitude of this single source in ASIA-AQv3 is significantly larger than the NO_x emissions allocated to known major sources like the Mae Moh power plant or even the Bangkok Metropolitan Region, within the same inventory. This large Lampang emission does not appear in the CAMS or the HTAP global inventories (Fig. S6), and there are no known large power plants or industrial facilities in Lampang city that would justify such a high emission. Notably, the total NO_x emissions for Thailand in ASIA-AQv3 are of the same order as those in CAMS and HTAP, implying that the Lampang point source significantly contributes to total NO_x emissions in Thailand. In other inventories, by contrast, the Bangkok Metropolitan Region and the Mae Moh power plant are the main NO_x contributors. This suggests that ASIA-AQv3's Lampang point source might be too high for NO_x emissions in Thailand.

Our inversion results indicate that ASIA-AQv3 indirectly affects the top-down emission estimates. The IFDMB posterior NO_x emissions for North Thailand appear to be influenced by the Lampang point source (Fig. S7). Because our WRF-Chem model and GEMS data are aggregated onto a coarse grid ($45 \times 45 \text{ km}^2$), any extreme point source signal in the GEMS observations tends to be spatially spread out over the surrounding area. As a result, the influence of the Lampang point source can be seen in the posterior emission field over a broad area of North Thailand. However, since we did not use ASIA-AQv3 directly to adjust the emissions, the posterior emissions do not simply replicate the ASIA-AQv3 spatial pattern.

Furthermore, the GEMS retrieval reported high RMSE for NO_2 VCDs in the area around this Lampang point source. In future applications of GEMS data, we recommend filtering the satellite data not only by the standard quality flags but also by excluding data points with abnormally high errors. This is because the RMSE (order of 10^{13}) in GEMS version 3 is relatively low compared to the NO_2 VCDs (order of 10^{15}). Consequently, the weighting by σ in the cost function is insufficient to handle this uncertainty. Applying such an error filter would help remove spurious signals and improve the robustness of emission updates derived from GEMS satellite observations.

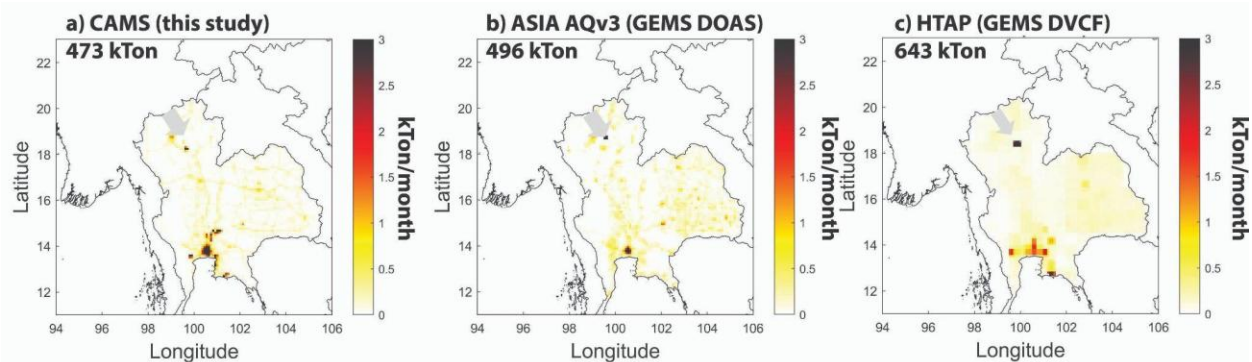


Figure S6: The comparison of NO_x emissions during September 2023 from three different inventories: a) CAMS, b) ASIA AQv3, and c) HTAP. The arrow points to the high NO₂ VCDs in GEMS DOAS.

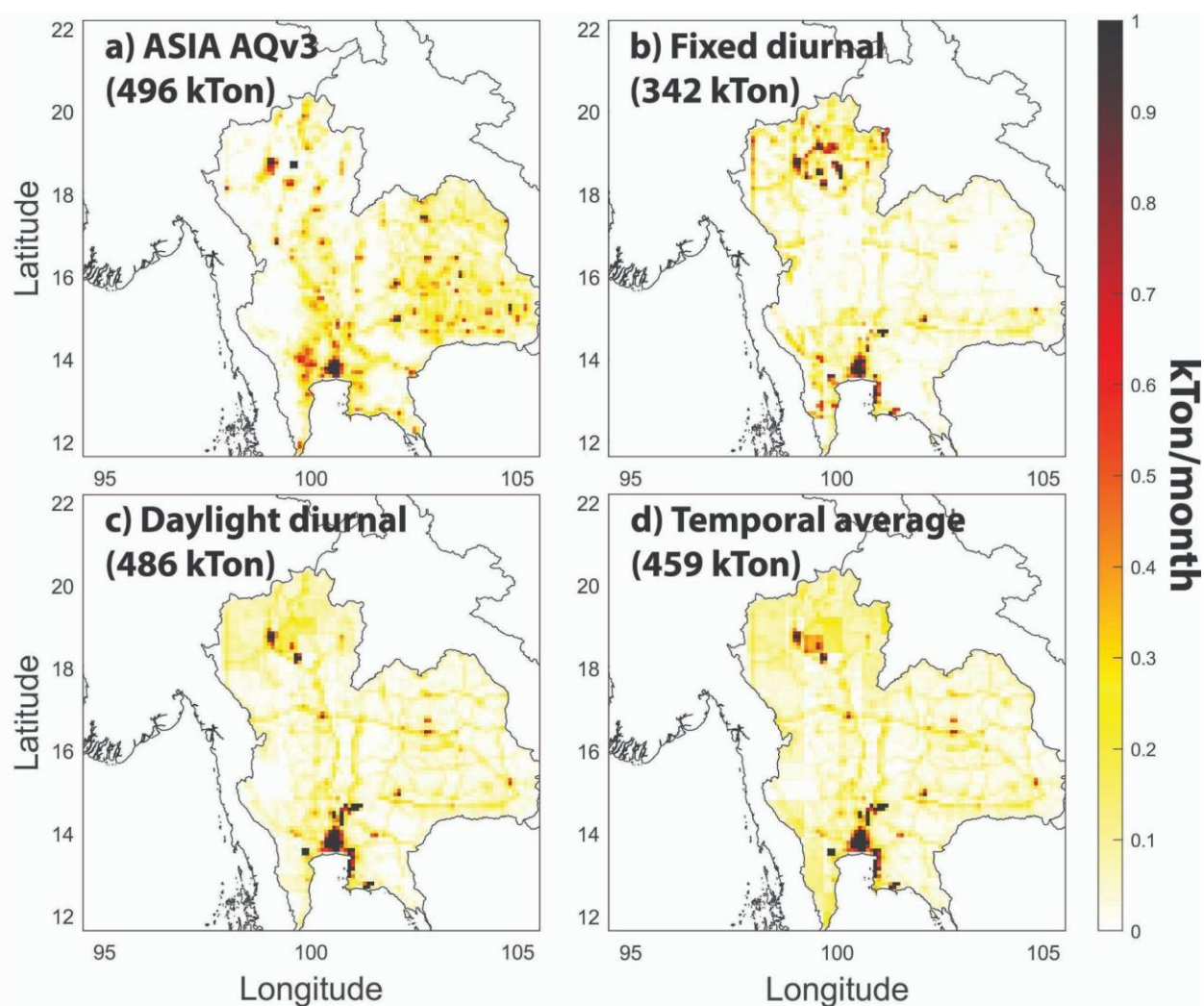


Figure S7: The comparison of NO_x emissions between ASIA AQv3 and our posterior emissions in September 2023.

S4 Moving Average for IFDDB daylight diurnal update

Because our IFDDB method only updates emissions during the GEMS observation window and reverts to prior emissions outside this period, it generates an unnatural, discontinuous emission profile. In this section, we test the ability of a smoothing technique to alleviate discontinuities caused by the temporal limitations of GEMS data. We apply the smoothing technique exclusively to the posterior emissions of the IFDDB with the daylight diurnal update. Our comparison focuses primarily on PCD stations, as the emissions within the GEMS observation window remain identical; therefore, the smoothing has no impact on TROPOMI comparisons and only a minor impact on GEMS at 08:00.

Specifically, we applied a 5-point moving average to bridge the GEMS daylight observation window (08:00–15:00) and the nighttime prior. The data within the GEMS daylight window remains unsmoothed to fully preserve the observational updates. For the transition periods (02:00–07:00 and 16:00–21:00), we applied the moving average using a distance-decaying weight (equation below).

$$E_{t,mov} = \frac{E_{t-2} + E_{t-1} + E_t + E_{t+1} + E_{t+2}}{5} \quad (S2)$$

$$E_{t,smooth} = (w - 1)E_t + wE_{t,mov} \quad (S3)$$

where $E_{t,mov}$ is the moving average of posterior emission at time t (E_t). The smoothed posterior emissions at time t ($E_{t,smooth}$) are calculated based on a weighted average between E_t and $E_{t,mov}$. During the GEMS observation window, the weight (w) is manually set to zero to preserve information from observation. For the 3 hours adjacent to the GEMS window (e.g., 05:00, 06:00, 07:00), we set w to 1 for fully smooth as shown in Figure S8. This is because these points are the transition zone between the GEMS observation window and the no-observation zone. As the time moves further from the GEMS window, this weight is reduced by 25% per hour until it reaches 0 (Figure S8). There is no GEMS information in this zone, but the smoothing is applied for smoothness.

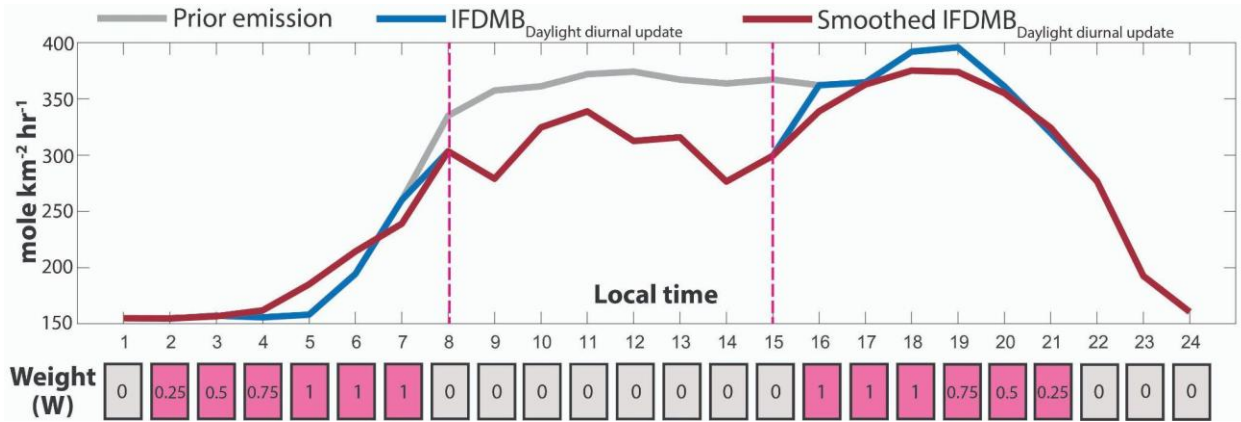


Figure S8: Hourly NO_x emissions in Bangkok comparing IFDDB with and without 5-point moving average using a distance-decaying weight: prior emission (gray), IFDDB with daylight diurnal update (blue), Smoothed IFDDB with daylight diurnal update (Red). The

pink line represents the GEMS window for Thailand, and the boxes show the weight of each time for Eq. (S3).

For example, the data at 04:00 uses the moving average of 02:00 to 06:00 with a weight of 0.75. The data at 03:00 uses a weight of 0.50, and at 02:00 uses a weight of 0.25. The evening transition (16:00–21:00) follows the exact same tapering logic. This smoothly relaxes the posterior daytime emissions back to the prior at night, successfully removing the unnatural discontinuity. However, the emission at 05:00 will be slightly oversmoothed. As this point is in the twilight zone, the 5-point moving average is used to smooth the emissions of daytime and nighttime to calculate smoothed posterior emissions.

When comparing against PCD stations, we found that while the smoothing method has a negligible impact on absolute NO_x concentrations (Fig S9), it produces a much smoother and more physically realistic NO_x concentration profile (Fig S10). Specifically, the smoothed IFDMB with the daylight diurnal update prevents the abrupt spikes and drops in concentration during the transition periods (15:00–17:00 and 06:00–08:00) caused by the sudden reversion of posterior emissions to prior emissions. This demonstrates the benefit of the smoothing technique. Even though the absolute concentrations do not change significantly, the resulting profile is much more realistic than the non-smoothed version. Future work employing more advanced smoothing techniques and incorporating nighttime observations will further refine the accuracy of these emission profiles.

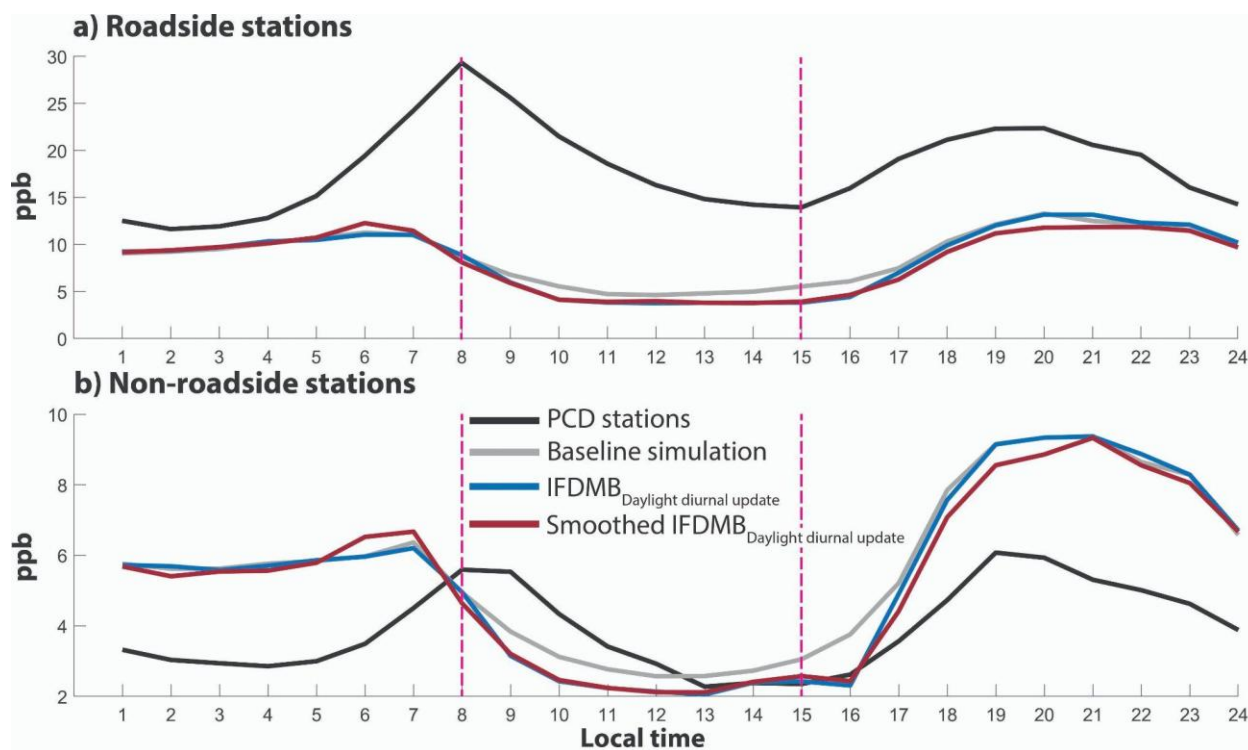


Figure S9: The monthly average diurnal profile of surface NO_x mixing ratios from PCD stations (black) and WRF-Chem: the baseline simulation (gray), IFDMB with daylight diurnal update (blue), and smoothed IFDMB with daylight diurnal update (red). PCD

stations are classified into roadside stations (a) and non-roadside stations (b). The pink lines mark the GEMS overpass times in Thailand from 08:00 - 15:00 local time.

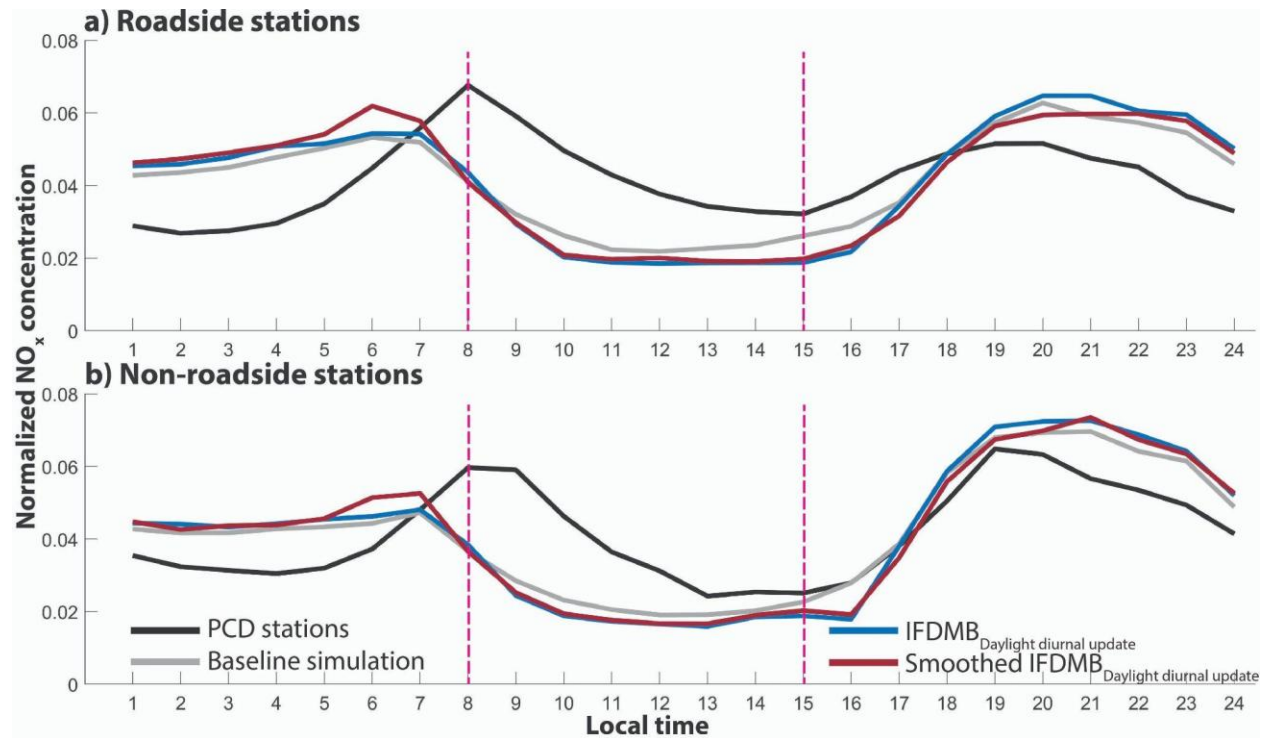


Figure S10: The diurnal variability of normalized NO_x concentrations from PCD stations (black) and WRF-Chem: the baseline simulation (gray), IFDMB with daylight diurnal update (blue), and smoothed IFDMB with daylight diurnal update (red). PCD stations are classified into roadside stations (a) and non-roadside stations (b). The pink line is the GEMS scanning time in Thailand from 08:00 - 15:00 local time.

S5 The analysis for the degree of grid cell aggregation

In this section, we quantify the impact of grid cell aggregation on the IFDMB with the fixed diurnal update scheme. We compare three degrees of aggregation: 5×5 grid cells (45×45 km, the original IFDMB configuration used in the main text), 3×3 grid cells (27×27 km), and no aggregation (1×1 , or 9×9 km).

For the comparison with GEMS (Table S2), no aggregation yields the lowest NMB and NRMSE among the three degrees of aggregation. This is because no aggregation overfits the GEMS satellite data by minimizing resolution error compared to others. However, the correlation coefficient for no aggregation is the lowest of the three schemes over Thailand, which we attribute to smearing error that leads to an underestimated adjustment in source areas. The 5-grid and 3-grid aggregations perform comparably in NRMSE and correlation, although the 5-grid aggregation gives a higher NMB over Thailand. Over Bangkok Metropolitan Region, no aggregation underestimates NO_2 VCD, whereas both the 5-grid and 3-grid aggregations overestimate it. This suggests that in large NO_2 emission hotspots such as Bangkok or industrial estates, smearing error is significant. The grid cell aggregation helps prevent the underestimation of emission adjustments in these areas. The 5-grid shows slightly lower NMB and NRMSE in Bangkok Metropolitan Region compared to 3-grid. However, no aggregation still provides the lowest NMB and NRMSE in this area.

For the comparison with TROPOMI (Table S3), no aggregation again provides the lowest NMB and NRMSE across Thailand as a whole, though the correlation coefficients are comparable across all three aggregation schemes, at 0.61 (5-grid), 0.68 (3-grid), and 0.63 (no aggregation). Because GEMS retrievals carry greater uncertainty over North Thailand due to the a priori profile used in the retrieval, we focus our comparison on the Bangkok Metropolitan Region instead. In this area, no aggregation gives the lowest NMB and NRMSE but also the lowest correlation coefficient. While the 5-grid and 3-grid aggregations perform comparably in NMB and NRMSE, the 5-grid aggregation achieves the highest correlation in this area. We interpret this as follows: in dense NO_2 source areas such as Bangkok Metropolitan Region or industrial estates, where NO_2 is transported further from its sources, smearing error is more important for the satellite data, so grid cell aggregation is beneficial because it reduces this error by preventing under-adjustment. In rural or agricultural areas, where NO_2 does not travel far from its source, resolution error dominates instead, so aggregation increases the bias and error in these areas. The 5-grid aggregation shows the highest correlation over the Bangkok Metropolitan Region in particular, likely reflecting the higher wind speeds typical of this coastal area.

For the comparison with PCD stations, all three aggregation schemes yield nearly identical results at roadside stations, with no significant difference in absolute concentration (Figs S11 and S12). This is because GEMS has a resolution of approximately 5 km and our WRF-Chem simulation a resolution of 9 km, neither of which is fine enough to resolve near-road transport of emissions. At non-roadside stations, however, the 5-grid aggregation performs better than the 3-grid and no aggregation (Fig S11). Because most non-roadside stations are located near

urban and industrial areas, they are more strongly affected by smearing error associated with the transport of NO₂ away from its sources, and therefore benefit most from the 5-grid aggregation.

Table S2: Monthly statistical evaluation between GEMS NO₂ VCD and WRF-Chem for different degrees of grid cell aggregation using IFDDB with fixed update schemes for Thailand and BMR.

Thailand	NMB	NRMSE	R
Baseline simulation	83.52	151.94	0.65
5 grids	31.90	56.25	0.84
3 grids	16.26	53.61	0.84
no aggregation	6.80	39.74	0.81
BMR	NMB	NRMSE	R
Baseline simulation	278.67	336.69	0.96
5 grids	29.46	51.93	0.95
3 grids	40.02	66.99	0.96
no aggregation	-12.77	43.01	0.96

Table S3: Monthly statistical evaluation between TROPOMI NO₂ VCD and WRF-Chem for different degrees of grid cell aggregation using IFDDB with fixed update schemes for Thailand and BMR.

Thailand	NMB	NRMSE	R
Baseline simulation	0.80	1.21	0.79
5 grids	0.60	0.86	0.61
3 grids	0.45	0.71	0.68
no aggregation	0.39	0.66	0.63
BMR	NMB	NRMSE	R
Baseline simulation	3.46	5.26	0.90
5 grids	0.59	1.25	0.93
3 grids	0.55	1.47	0.86
no aggregation	-0.07	0.93	0.80

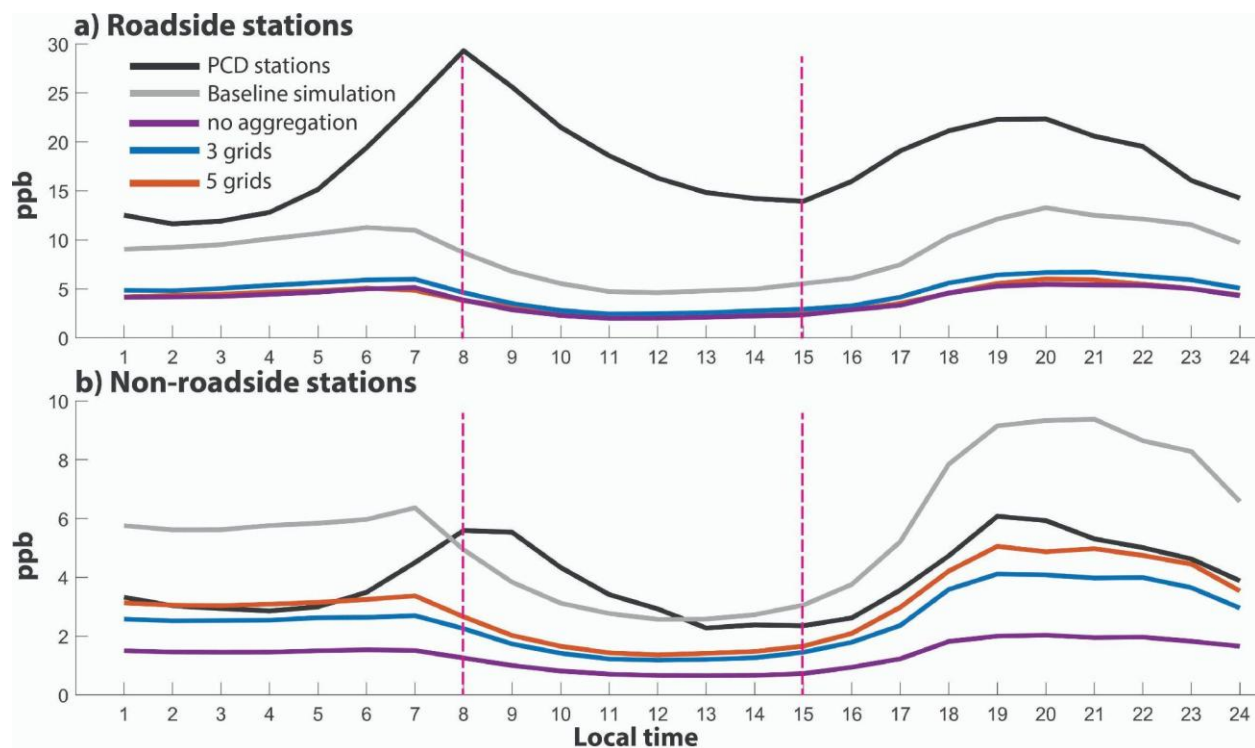


Figure S11: The monthly average diurnal profile of surface NO_x mixing ratios from PCD stations (black) and WRF-Chem: the baseline simulation (gray), no aggregation (purple), 3-grid aggregation (blue), and 5-grid aggregation (orange). PCD stations are classified into roadside stations (a) and non-roadside stations (b). The pink lines mark the GEMS overpass times in Thailand from 08:00 - 15:00 local time.

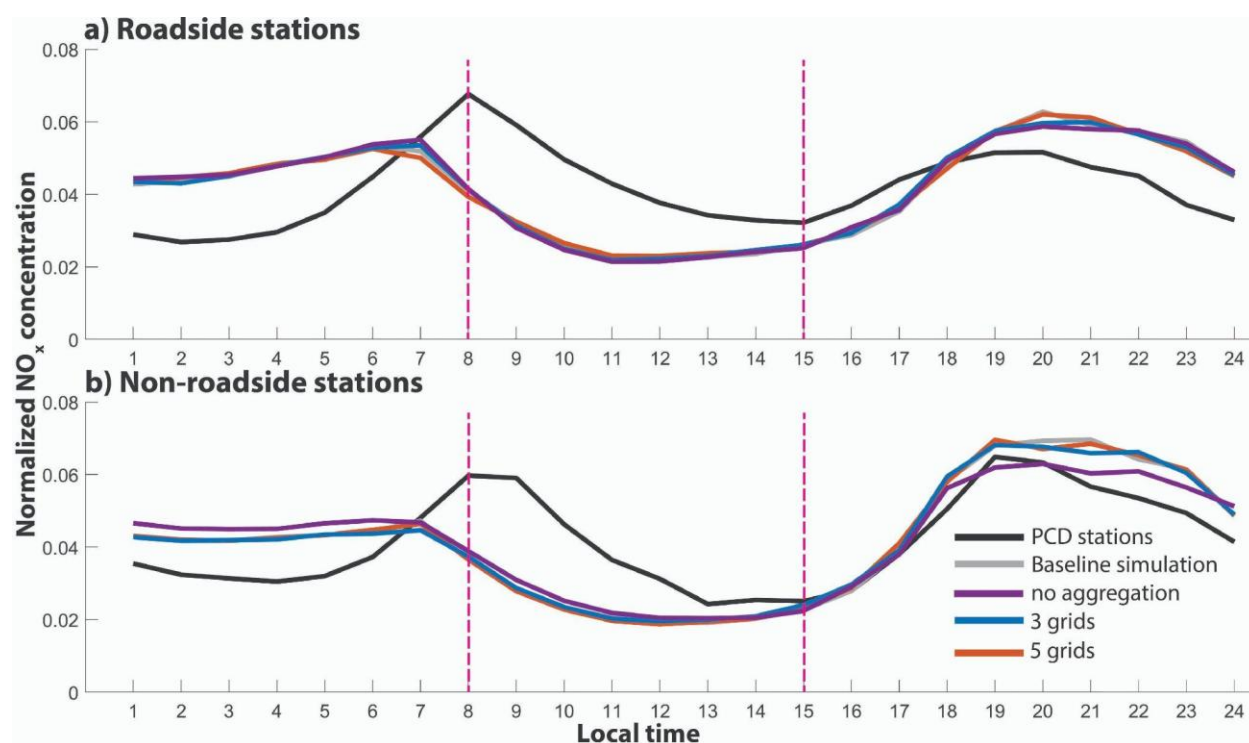


Figure S12: The diurnal variability of normalized NO_x concentrations from PCD stations (black) and WRF-Chem: the baseline simulation (gray), no aggregation (purple), 3-grid aggregation (blue), and 5-grid aggregation (orange). PCD stations are classified into roadside stations (a) and non-roadside stations (b). The pink line is the GEMS scanning time in Thailand from 08:00 - 15:00 local time.

S6 Averaging kernel

Retrieving NO₂ from ultraviolet and visible satellite observations relies on three stages (Cooper et al., 2020): calculating a slant column, removing the stratospheric contribution, and converting the slant column into a vertical column density (VCD) using an air mass factor (AMF). AMFs are calculated using a radiative transfer model and are dependent on viewing geometry, surface reflectance, and cloud cover (Eskes and Boersma, 2003). Furthermore, these calculations require an a priori estimate of the NO₂ vertical profile (Cooper et al., 2020; Eskes and Boersma, 2003). Uncertainties in the AMF and the a priori information used in the retrieval algorithm can degrade the quality of model–satellite comparisons beyond the inherent errors of either the model or the measurements alone (Boersma et al., 2007). To mitigate these errors, an averaging kernel (AK) is commonly applied to the model output. The AK accounts for the altitude-dependent sensitivity of the satellite, providing a consistent mathematical relationship between the retrieved quantities and the true tracer profile (Eskes and Boersma, 2003).

To account for the sensitivity of the satellite instrument, the AK is applied to the interpolated model profile using the following equation (Eskes and Boersma, 2003):

$$\Omega_{WRF_{satellite}} = \sum_{i=1}^{i=tropopause} AK_i \Omega_{WRF,i} \quad (S4)$$

where $\Omega_{WRF_{satellite}}$ is the tropospheric vertical column density (VCD) accounting for satellite sensitivity. AK_i is the averaging kernel from satellite data at layer i . $\Omega_{WRF,i}$ is the simulated WRF-Chem NO₂ vertical column density at layer i , after being mapped to the satellite's vertical pressure grid.

The AK is mathematically derived from the AMF and scattering weights. Because these values depend on the specific retrieval algorithm and the assumed a priori profiles, the resulting AKs differ inherently between GEMS and TROPOMI. To evaluate the relative sensitivities of these satellites, the monthly mean surface of the first layer ($i = 1$) AKs for both GEMS and TROPOMI over Thailand are compared in Fig. S13. The monthly mean AK_{GEMS} aggregates all daylight diurnal data points (00:45–04:45 UTC and 06:45–07:45 UTC), whereas the monthly mean $AK_{TROPOMI}$ is based solely on the single daily overpass at approximately 06:30 UTC. While the AK spatial distributions across Central Thailand and the Bangkok Metropolitan Region are similar between the two satellites, distinct disparities emerge in Northern Thailand. TROPOMI exhibits elevated surface AK over Chiang Mai and the Mae Moh power plant, whereas GEMS shows higher AK over the broader Lampang area. These regional discrepancies are largely attributable to uncertainties in the a priori profiles utilized by the GEMS retrieval algorithm. Additionally, significant AK variations are shown in Rayong, a major industrial estate containing extensive harbor and power plants. $AK_{TROPOMI}$ demonstrates high AK along the coastal areas, capturing emissions from shipping, coastal power plants, and near-shore industrial zones. Conversely, AK_{GEMS} exhibits its highest values in inland industries and power plants. Ultimately, these spatial discrepancies confirm that the retrieval algorithms of the two satellites possess distinctly different sensitivities to surface-level NO₂ concentrations.

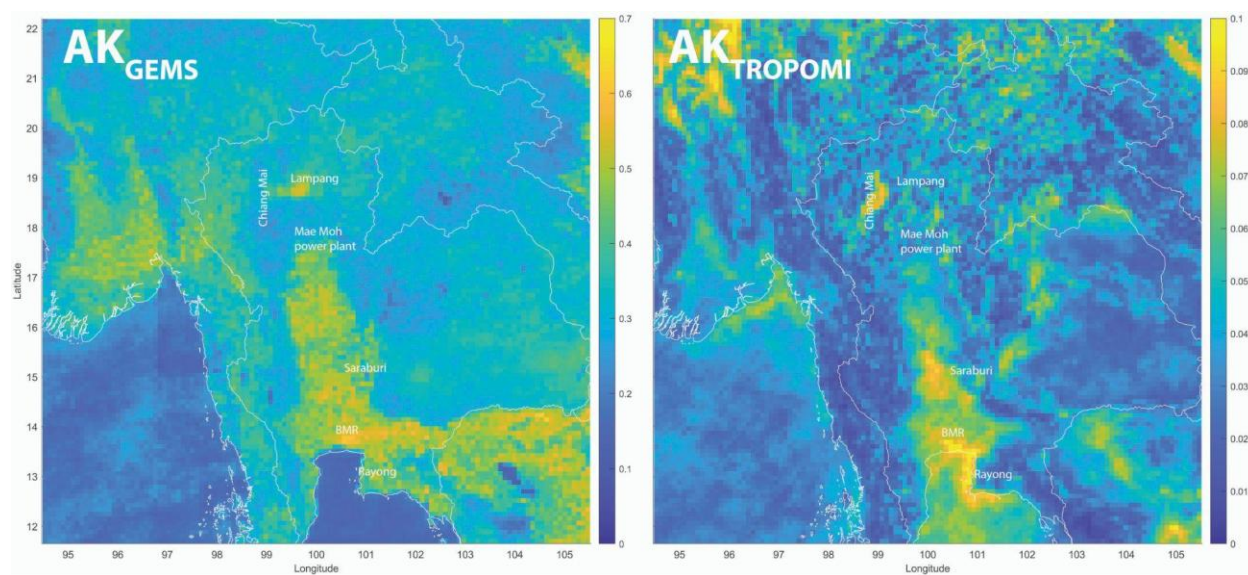


Figure S13: The monthly mean averaging kernel for GEMS and TROPOMI at the surface (1st layer).

Supplement figures

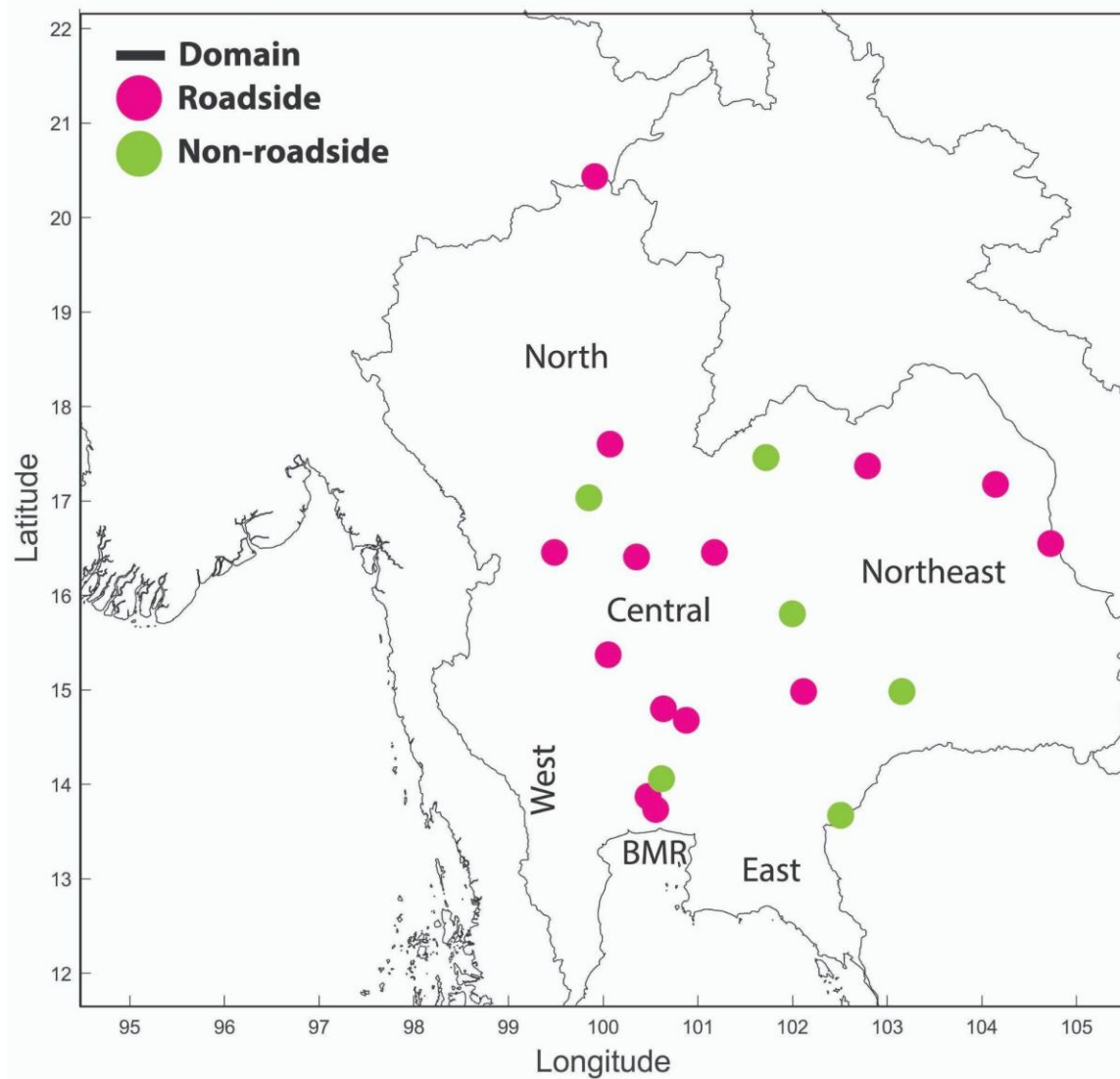


Figure S14: Model domain and location of PCD stations, classified as roadside (pink) and non-roadside (green). BMR is the Bangkok Metropolitan Region.

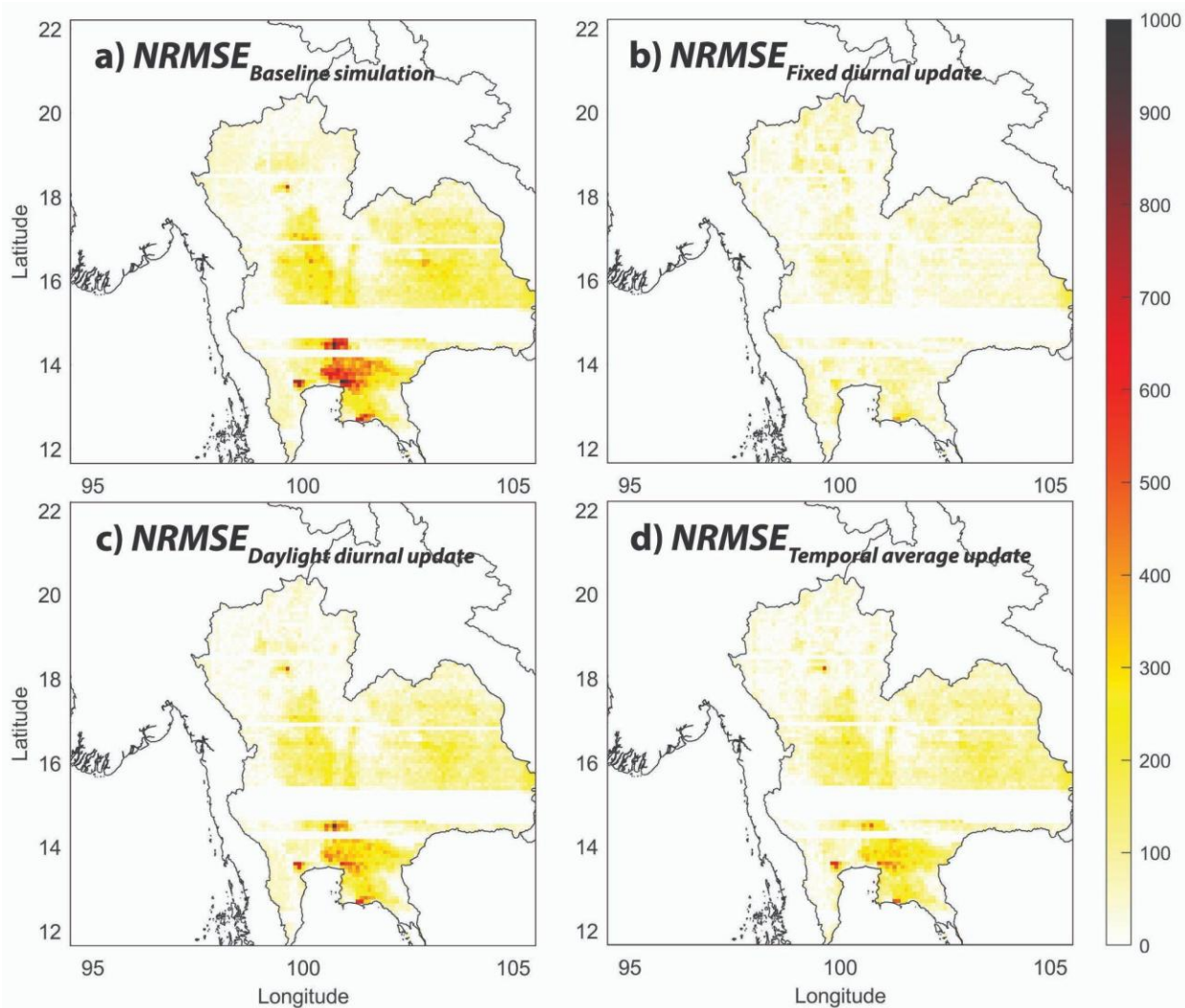


Figure S15: Normalized Root Mean Square (NRMSE) of NO₂ VCD between WRF-Chem and GEMS of IFDMB for September 2023 using (a) prior emissions, and emissions updated by (b) fixed diurnal, (c) daylight diurnal, and (d) temporal average update schemes.

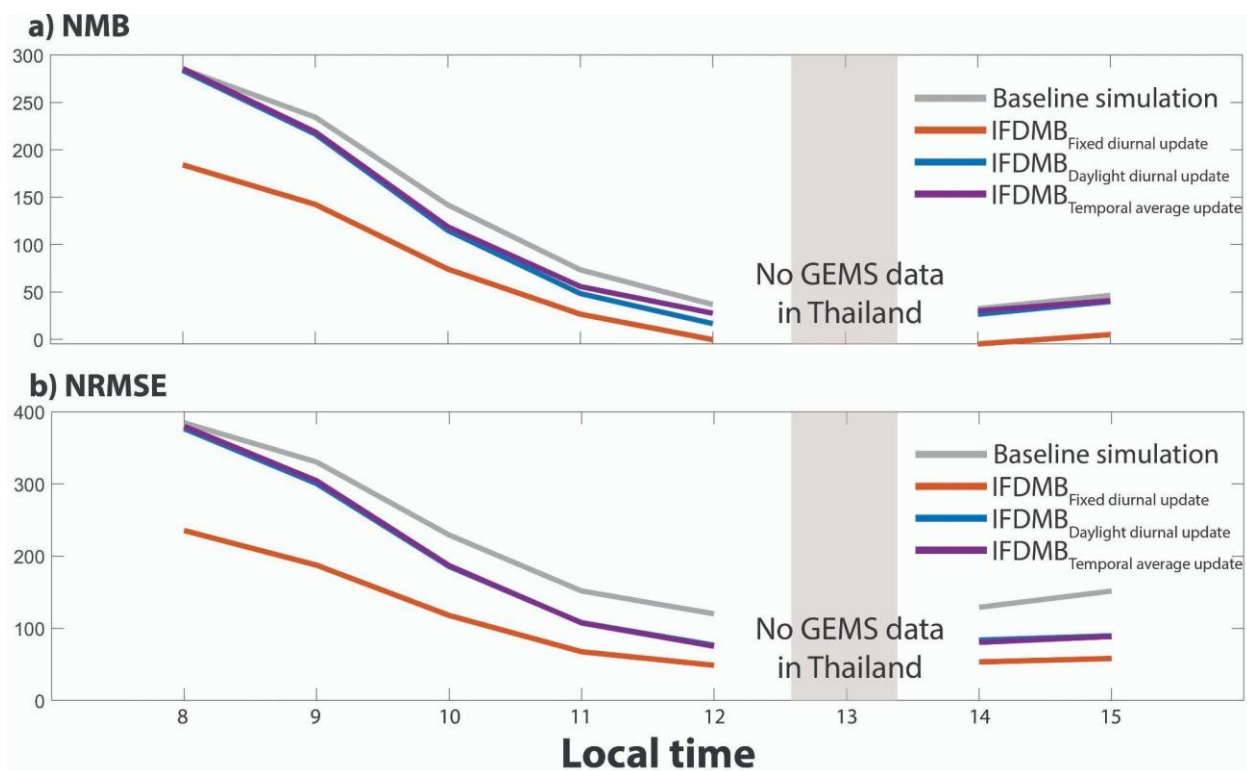


Figure S16: Monthly statistical evaluation between WRF-Chem and GEMS NO₂ VCD of each hour averaged from Thailand grid cells.

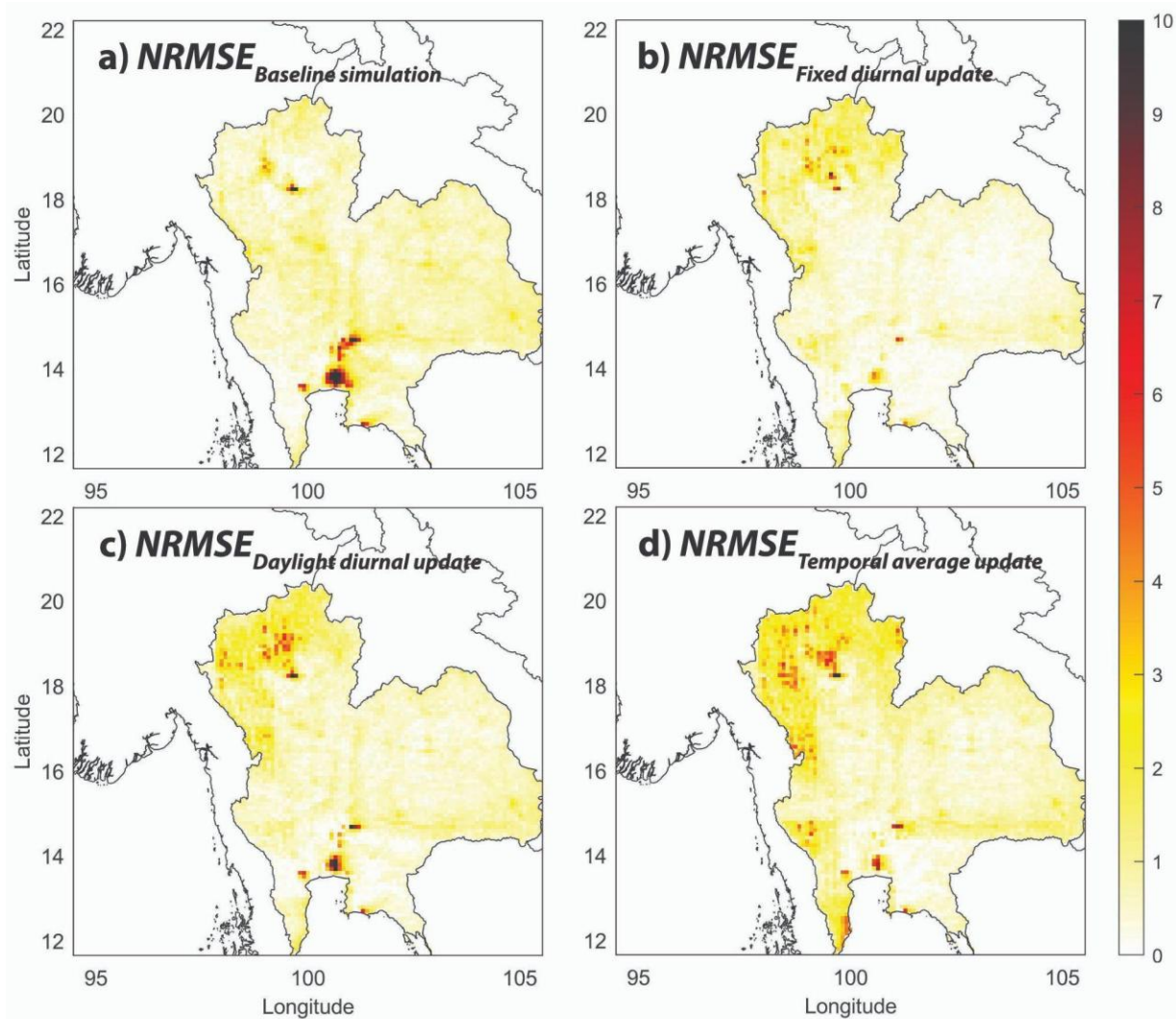


Figure S17: NRMSE of NO₂ VCD between WRF-Chem and TROPOMI of IFDMB for September 2023 using (a) prior emissions, and emissions updated by (b) fixed diurnal, (c) daylight diurnal, and (d) temporal average update schemes.

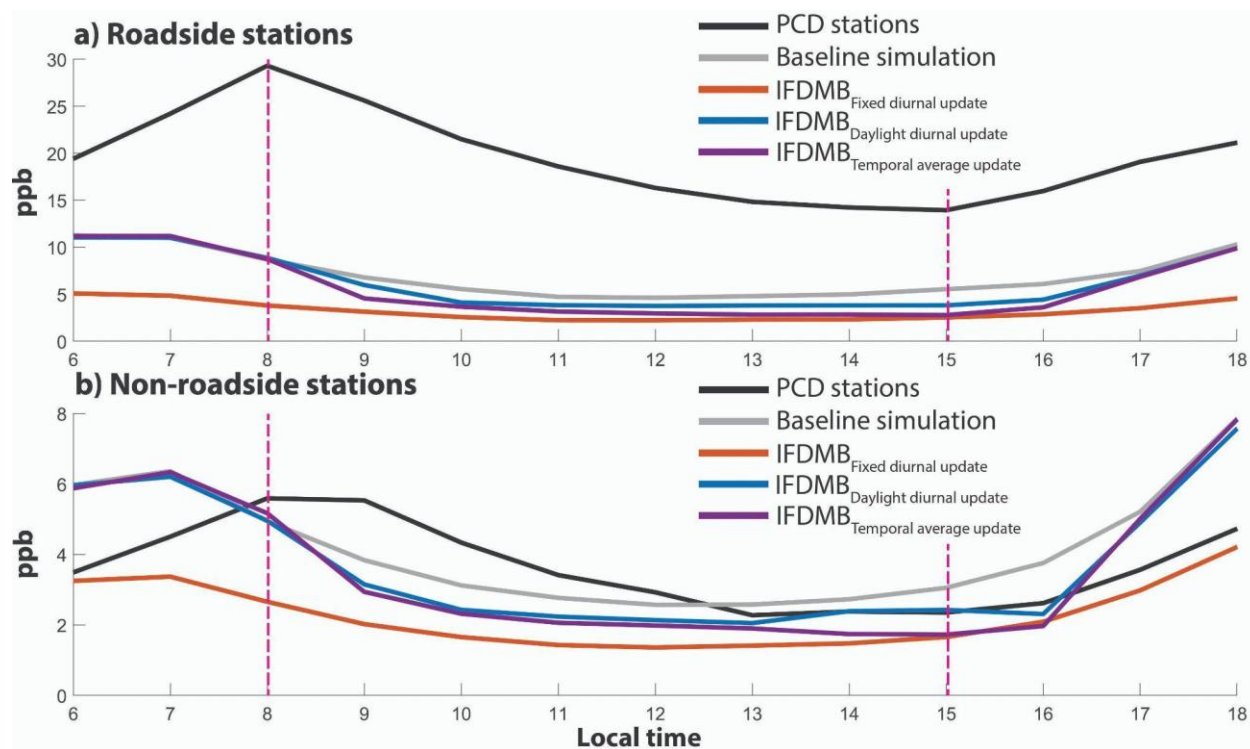


Figure S18: The monthly average diurnal profile of surface NO_x mixing ratios from PCD stations (black) and WRF-Chem: the baseline simulation (gray), IFDMB with fixed diurnal update (orange), IFDMB with daylight diurnal update (blue), and IFDMB with temporal average update (purple). PCD stations are classified into roadside stations (a) and non-roadside stations (b). The pink lines mark the GEMS overpass times in Thailand from 08:00 - 15:00 local time.

Supplement table

Table S4: Monthly statistical evaluation between TROPOMI NO₂ VCD and WRF-Chem for IFDMB with different update schemes for Thailand and BMR.

Thailand	NMB (unitless)	NRMSE (unitless)	<i>R</i>
Baseline simulation	0.80	1.21	0.79
IFDMB _{Fixed diurnal}	0.60	0.86	0.61
IFDMB _{Daylight diurnal}	0.84	1.19	0.63
IFDMB _{Temporal average}	1.05	1.41	0.47
BMR	NMB (unitless)	NRMSE (unitless)	<i>R</i>
Baseline simulation	3.46	5.26	0.90
IFDMB _{Fixed diurnal}	0.59	1.25	0.93
IFDMB _{Daylight diurnal}	1.58	3.22	0.88
IFDMB _{Temporal average}	0.66	2.05	0.82

References

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO₂ from OMI, *Atmos Chem Phys*, 2007.

Cooper, M. J., Martin, R. V., Henze, D. K., and Jones, D. B. A.: Effects of a priori profile shape assumptions on comparisons between satellite NO₂ columns and model simulations, *Atmos. Chem. Phys.*, 20, 7231–7241, <https://doi.org/10.5194/acp-20-7231-2020>, 2020.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, *Atmos Chem Phys*, 2003.

Huang, X., Yang, K., Kondragunta, S., Wei, Z., Valin, L., Szykman, J., and Goldberg, M.: NO₂ retrievals from NOAA-20 OMPS: Algorithm, evaluation, and observations of drastic changes during COVID-19, *Atmos. Environ.*, 290, 119367, <https://doi.org/10.1016/j.atmosenv.2022.119367>, 2022.

Park, J., Choi, Y., Jung, J., Lee, K., and Yeganeh, A. K.: First top-down diurnal adjustment to NO_x emissions inventory in Asia informed by the Geostationary Environment Monitoring Spectrometer (GEMS) tropospheric NO₂ columns, *Sci. Rep.*, 14, 24338, <https://doi.org/10.1038/s41598-024-76223-1>, 2024.

Yang, K., Carn, S. A., Ge, C., Wang, J., and Dickerson, R. R.: Advancing measurements of tropospheric NO₂ from space: New algorithm and first global results from OMPS, *Geophys. Res. Lett.*, 41, 4777–4786, <https://doi.org/10.1002/2014GL060136>, 2014.

Yang, K., Kondragunta, S., and Wei, Z.: The NO₂ Algorithm for GeoXO-ACX and Application to GEMS and TEMPO, <https://doi.org/10.22541/essoar.170365312.22362491/v1>, 2023.